

Is there such a thing as free government data?

Original

Is there such a thing as free government data? / Morando, Federico; Iemma, Raimondo; Basso, Simone. - In: INTERNET POLICY REVIEW. - ISSN 2197-6775. - ELETTRONICO. - (2013).

Availability:

This version is available at: 11583/2519717 since:

Publisher:

Alexander von Humboldt Institute for Internet and Society

Published

DOI:

Terms of use:

openAccess

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

(Article begins on next page)

Is there such a thing as free government data?

Federico Morando and Raimondo Iemma and Simone Basso

The recently-amended European public sector information (PSI) directive (Directive 2013/37/EU, PDF, hereinafter “the directive”) rests on the assumption that “[d]ocuments produced by public sector bodies of the Member States constitute a vast, diverse and valuable pool of resources that can benefit the knowledge economy” (recital 1).

More specifically, European policy-makers submit that “[o]pen data policies which encourage the wide availability and re-use of public sector information for private or commercial purposes, with minimal or no legal, technical or financial constraints [...] can play an important role in kick-starting the development of new services [...], stimulate economic growth and promote social engagement” (recital 3).

Therefore, to keep financial constraints on re-use as low as possible, the directive provides that, “where charges are made by public sector bodies for the re-use of documents, those charges should in principle be limited to the marginal costs.” In practice, this should imply that most (natively digital) government data are free to re-use for any (lawful) purpose.

This article provides a brief review of the public sector information pricing issues. It then discusses the terms under which the ongoing consultation on the implementation guidelines of the PSI directive addresses pricing. In particular, this article discusses the calculation criteria for marginal costs.

The simple economics of PSI charging

Digital goods hold well known features: their creation entails high fixed costs, while reproducing them is almost costless. As a consequence, charging for them is typically tricky. This issue has been thoroughly debated by economists, who perhaps got inspired by the so-called ‘marginal cost controversy’ (PDF), dating back to 1946, which involved Ronald Coase and Harold Hotelling, debating the optimal charging principles for public goods, and in particular whether marginal cost pricing, or charges allowing to recoup also fixed costs (e.g., two-part tariffs), were to be assumed as more desirable in terms of overall welfare.

One should also consider at least two other features of digital PSI. First of all, it has great potential for re-use. In fact, governments collect and manage tremendous amounts of information, which is assumed to be complete and accurate, and which, in many cases, is the only possible source of the data that one might want to embed in a digital service (Pollock, 2009). Secondly, where data stems as an incidental by-product of the public task, the PSI production has been already funded through taxation (LAPSI position paper nr. 1, 2011, PDF).

Because of the two features mentioned above, and because PSI entails both supply and demand side economies of scale, several economists (e.g., Koski, 2011; Pollock, 2011), as well as empirical studies such as POPSIS, highlight the positive externalities, for example in terms of economic growth, generated by a wider circulation of PSI, also driven by charges equal to marginal costs (or even lower, i.e., zero charges). One should also consider that getting paid makes further types of costs arise, i.e., transaction costs. When transaction costs are higher than the marginal cost-based price, the public administration should make the PSI available free of charge.

Yet, there are also reasonable arguments against low charges. For instance, if a public agency is not the unique holder of a specific dataset, by giving the data away for free, the agency may be implementing predatory pricing. Moreover, the free of charge strategy is typically coupled with a best effort level of service, while (at least) for-profit re-uses arguably need high-quality data as an input.

In the experience deriving from policy support activities performed by the authors at regional and European

level, in the common practice, at least two charging approaches seem to coexist and to be applied to different segments, or 'low-end' and 'high-end' markets as identified in the POPSIS study (PDF). Public administrations currently make available small, previously undisclosed datasets at no charge. At the same time, national agencies extract profits from licensing access to databases of high interest, e.g., firm registries or geodata.

PSI charging: a brief history of a long debate

It could be argued that the Guidelines for improving the synergy between public and private sectors in the information market (PDF), promoted by the EC in 1989, represent the first step towards the definition of a European PSI policy. The Green paper on public sector information in the information society (PDF), published in 1999 by the European Commission, continued on this track, with an explicit focus on PSI. The Green Paper contained a review of the main issues at stake, including pricing. Also, the Green Paper argued that the optimal charging principles should strike a balance between allowing affordable access to everyone, fostering the exploitation potential of PSI, and ensuring fair competition.

Four years later, the European Commission issued the first piece of legislation addressing PSI, the European directive 2003/98 (PDF). In a nutshell, with respect to charging (if any), the 2003/98 directive allowed PSI holders to recoup collection, production, reproduction and dissemination costs, together with a reasonable return on investment. In 2010, the European Commission promoted a public consultation in view of the revision of the PSI directive: the vast majority of the respondents signalled that PSI re-use had not achieved its full potential in Europe; around 40% of the respondents agreed on the marginal cost principle (reproduction and dissemination) for PSI pricing, while 36% disagreed; in any case, 54% of the participants were in favour of tightening and/or making more clear charging rules.

In June 2013, the amended version of the European directive on the re-use of public sector information was issued, containing, amongst other changes, updated prescriptions on charging.

Charges in the amended EU PSI Directive

Article 6 of the amended PSI directive discusses the principles governing charging, which we summarise below.

The new default rule is that charges for the re-use of the PSI have an upper bound in the "marginal costs incurred for [the] reproduction, provision and dissemination" of government data (§1 of art. 6).

As an exception, the directive allows the public sector bodies (PSBs) to charge higher fees in cases in which they are "required [by the law or by administrative practices] to generate sufficient revenue to cover a substantial part of the costs relating to their collection, production, reproduction and dissemination" (§2). If they charge higher fees, the PSBs must set charges according to objective, transparent and verifiable criteria; moreover the total income from supplying and allowing re-use of documents must "not exceed the cost of collection, production, reproduction and dissemination, together with a reasonable return on investment" (§3).

Another exception to the default rule is that libraries, archives and museums (LAMs) are generally allowed to charge above marginal costs. Moreover, LAMs charges can also take into account the costs of "preservation and rights clearance" (§ 4).

The ongoing consultation: guidelines on charging

European Union member states are free to apply lower charges and, in particular, no charges at all. This freedom is consistent with the directive, which primarily aims at maximising the PSI re-use and its economic benefits, as well as with the principle of minimum harmonisation. Moreover, the principle of subsidiarity imposes that the criteria for charging above marginal costs are essentially left to member states (recital 25).

However (and as stated in recital 36), the Commission shall help the member states to implement the directive “in a consistent way by issuing guidelines, particularly on recommended standard licences, datasets and charging for the re-use of documents, after consulting interested parties.”

In the following paragraphs we focus on the ongoing consultation envisioned by the directive and, in particular, on its fourth section, which deals with the practical implementation of charges for the re-use of the PSI. By taking advantage of the questions as spelled in the consultation, we proceed to analyse the key open issues.

Calculating the marginal cost of public sector information

To implement the directive, an operational rule to calculate the marginal cost of “reproduction, provision and dissemination” is needed. To this end, the consultation asks the respondents whether the following cost items should contribute to the calculation of marginal costs: telecommunications costs, customer service, duplication, software licensing, database modification(s) for dissemination, hardware enhancements for dissemination (capacity, ports), value-added (activities) for dissemination (software enhancements, advertising), database development(s), hardware, data creation/collection, data maintenance, and archiving.

The consultation allows the responders to choose between the four following answers: always, until amortised, never, and no opinion.

The standard definition of marginal cost as the change in total costs that arises when the quantity produced is increased by one unit (i.e., the cost of producing one more unit of a good) suggests the following answers:

1. duplication costs always contribute to the marginal costs. In practice, however, in a digital environment, the duplication cost is zero (except when the original data is in analog format and must be digitised);
2. telecommunications and customer service costs could or could not be marginal costs; in principle, one should answer with the no opinion option and should use the open answer option to provide the following explanation.

First and foremost: some marginal “telecommunications costs” do exist. For instance, the cost of adding network capacity to satisfy a certain request is a marginal cost; to recoup this cost, one can, for example, implement a “capacity charge” that captures the amount of capacity consumed by a user. Similarly, the customer-service costs generated by a user can be charged to her/him, e.g., by using a premium-rate telephone number (i.e., the 900- or 199- numbers, depending on the country). In other words, the marginal telecommunication cost directly generated by the *i*th re-user should be charged to the *i*th re-user only.

However, not all telecommunication services allow their owners to charge their users in a simple way. A premium-rate telephone number, in fact, allows its owner to bill the user that makes the phone call, but the internet lacks a money-routing protocol and lacks a per-flow charging mechanism.

Therefore, when the costs of an internet-based service are significant, a reasonable answer may be until amortised: in practice, instead of charging the *i*th re-user the cost he/she generates, one estimates the expected number of re-users, *N*, and charges all re-users $1/N$ of the cost of allowing re-use (or a better/easier re-use). However, this approach is not theoretically compatible with the new directive, because it does not consider the actual “marginal cost” of re-use, but what can be described as the “average incremental cost” of allowing re-use;

3. the sub-set of cost items “for dissemination” (i.e., database modifications, hardware enhancements such as capacity and ports, value-added activities such as software enhancements and advertising) contributes to the average incremental cost of allowing re-use as well. Therefore, until amortised is again a reasonable, although theoretically incorrect, answer;

4. software licensing could be comprised in the previous point, however, as a matter of policy, this could (and

possibly should) be avoided: the rationale of never allowing to recoup licensing costs is that every needed activity in this domain can be performed with open source software at no charge (and at least some member states may want to encourage this approach);

5. database development(s), hardware, data creation/collection, data maintenance, and archiving should never be considered, as they are typical examples of fixed costs, which are sunk at the moment of making PSI accessible and re-usable. An additional reason not to charge these costs on PSI re-users is the following: doing otherwise would create an incentive for the PSBs to charge on the re-users costs, which are actually related with the overall ICT management of the PSBs themselves.

In conclusion, there are several cases in which the marginal cost approach, if strictly implemented, would imply that just the *i*th user (or, in certain cases, the first user) would pay a high fee, with users from *i*+1 onward receiving the improved service for free, if no further marginal costs are generated (e.g., user *i* requires some data to be published in a new format: a conversion tool is developed and paid by her/him, while the rest of the users can get the new format for free)¹. It may, however, be reasonable to treat these cases differently, guessing the total expected number of users (and/or shifting the costs on the next fiscal year) and charging on each of them pro rata: this is our understanding of the until amortised option offered in the consultation.

Special cases: full cost recovery scenarios

Article 6 of the PSI directive provides that, where full cost recovery is allowed, the total income from supplying and allowing the re-use of PSI “shall not exceed the cost of collection, production, reproduction and dissemination, together with a reasonable return on investment.” Accordingly, the consultation asks which of the following costs may be included in the calculation of fees for re-use: overheads, non-incremental database development costs, non-incremental hardware costs, data maintenance.

Considering the generic language of the directive and its permissive rationale, all these costs could possibly be considered (the other available answers being always, never and no opinion). That said, the consultation may arguably be criticised for its lack of precision in defining costs such as “non-incremental hardware costs”².

A related question concerns a definition of a “reasonable return on investment”. The consultation investigates what percentage above the fixed interest rate on the main refinancing operations set by the ECB (currently 0.5%) should be considered “reasonable”. From the economic point of view, one can set a “reasonable” return on investment by looking at the typical return on investment of a private player in a comparable, competitive market. However, private players typically demand a higher return on investment, considering, e.g., the risk of going bankrupt. Because PSBs do not typically go bankrupt, we submit that, intuitively, a moderate 2-5% premium over the main refinancing rate of the ECB could be provided as a reference point. It is however fair to consider this as a mere personal opinion of the authors.

Finally, another special case concerns libraries, museums and archives. Not only they are always free to charge more than marginal costs, they can also recoup additional cost elements, i.e., the cost of preservation and rights clearance. The consultation asks how these costs should be calculated, but the question appears to be partly tautological: any cost of preservation and rights clearance could arguably be recouped, possibly including the cost of digitisation itself and the cost of copyright searches (e.g., to assess whether a work is in the public domain). In this regard, i.e., in cases in which the re-used PSI consists of digitised public domain material, the most delicate point does not concern charges, but the fact that public domain material should arguably remain in the public domain as a matter of public policy. Therefore, as soon as one has a copy of a public domain piece of content, he or she should be free to use, re-use and share it as he or she sees fit. It is therefore very difficult to imagine on which basis it could be possible to charge any costs on the re-users, unless LAMs are allowed to contractually void the public domain status of these works of most of its meaning.

Final remarks and policy implications

Broadly speaking, the new European prescriptions concerning PSI charging may seem to be the result of an act of balance between the need of a wider and easier circulation of information, and the current budget constraints of public agencies. In practice, the radical option of marginal - and de facto zero - cost, attractive as it may be on paper, might be nothing more than a formal default option.

In several cases, a public administration might decide not to charge at all. This is indeed what has happened for all datasets made available through open-government data portals, even before the amended directive was issued. This approach may be appealing for PSBs also because, where charges are made, they have to be calculated following “objective, transparent and verifiable criteria” and doing so may involve some intricacies (and related costs).

Conversely, when it decides to charge, a public administration is always allowed to recoup the marginal cost of reproduction, provision and dissemination. But, as we discussed, it is complex to define this kind of cost, especially for internet-based services, and it is even harder to design charging policies based on it; in fact, the questions and (in particular) the answers of the PSI consultation on charging indirectly lead the responder to reason in terms of the average incremental cost of allowing re-use, instead of the marginal cost of reproduction, provision and dissemination. Unfortunately, as discussed in the section about the economics of PSI charging, it is marginal-cost pricing (or zero pricing) that is expected to lead to economically efficient results (while aiming at the recoupment of the average incremental cost of allowing re-use may lead to excessive fees).

Finally, under the current rules, it seems quite easy to take advantage of the allowed exceptions, especially for public agencies who so far relied on income deriving from PSI dissemination. Not by chance, databases with higher potential for commercial re-use are arguably held by those agencies, and feed mature re-use markets that usually hold strong barriers to entry.

References

Coase, R. (1946). The Marginal Cost Controversy. *Economica, New Series*, 13(51), 169-182.

Davies, R., (2008). Recommendation of the ePSIplus network to the EC review of the Directive on PSI re-use. *Paper presented at the first Communia conference*. Retrieved from URL <http://www.communia-project.eu/node/112>.

Koski, H. (2011). Does marginal cost pricing of public sector information spur firm growth? *ETLA Discussion Papers* no. 1260.

Pollock, R. (2009). The Economics of Public Sector Information. Retrieved from URL http://rufuspollock.org/papers/economics_of_psi.pdf

Pollock, R. (2011). Welfare gains from opening up Public Sector Information in the UK. University of Cambridge, undated, accessed 18 November 2013. Retrieved from URL http://rufuspollock.org/papers/psi_openness_gains.pdf

Ricolfi, M., Drexler, J., van Eechoud, M., Janssen, K., Maggolino, M.T., Morando, F., Sappa, C., Torremans, P., Uhlig, P., Iemma, R., de Vries, M. (2011). The 'principles governing charging' for re-use of public sector information. *Informatica e Diritto*, 1-2, 105-128.

Uhlig, P. (2009). *The Socioeconomic Effects of Public Sector Information on Digital Networks: An Analysis of Different Access and Reuse Policies: Workshop Summary*. Washington, D.C.: National Academies Press.

Footnotes

1. If this approach is chosen, notice that at least one exception should apply: no charging should be made in case the customisation request actually consists in a bug fixing, because this signalling activity should be subsidised, as a matter of policy, since it generates a public good (for all re-users and possibly for the PSB itself).
2. A Google search on "non-incremental hardware costs" just returns the text of the consultation, confirming the impression that this concept is far from being a commonly understood one.